

1. A method comprising:

transforming frames in a video sequence using a wavelet transform and motion information between frames to produce multiple sub-bands of coefficients; and

coding the coefficients of each sub-band independently.

2. A method as recited in claim 1, wherein the wavelet transform comprises a shape-adaptive discrete wavelet transform.

3. A method as recited in claim 1, wherein the transforming comprises performing a temporal 1-D wavelet transform along motion trajectories in a temporal direction.

4. A method as recited in claim 1, wherein the transforming comprises:  
performing a temporal wavelet transform on corresponding pixels in a video object along motion trajectories in a temporal direction to produce frames of temporal wavelet coefficients; and

performing a spatial wavelet transform on the frames of the temporal wavelet coefficients to produce multiple sub-bands of wavelet coefficients.

5. A method as recited in claim 1, wherein the coding produces multiple bitstreams, one for each sub-band, and further comprising forming a bitstream from the multiple bitstreams.

1        6. A method as recited in claim 1, wherein the coding comprises  
2 transposing selected sub-bands.

3  
4        7. A method as recited in claim 1, wherein the coding comprises coding  
5 the coefficients of each sub-band bit-plane by bit-plane using different coding  
6 primitives.

7  
8        8. A method as recited in claim 7, wherein the coding primitives  
9 comprise:

10        zero coding to code new information about a coefficient that is not yet  
11 significant in a previous bit-plane; and

12        sign coding to code a sign of the coefficient once the coefficient is deemed  
13 significant.

14  
15        9. A method as recited in claim 7, wherein the coding primitives  
16 comprise:

17        zero coding to code new information about a coefficient that is not yet  
18 significant in a previous bit-plane;

19        sign coding to code a sign of the coefficient once the coefficient is deemed  
20 significant; and

21        magnitude refinement to code new information of a coefficient that has  
22 already become significant in the previous bit-plane.

10. A method as recited in claim 1, wherein the coding comprises assigning contexts to the coefficients of each sub-band based on numbers of significant neighboring samples.

11. A method as recited in claim 10, wherein the sub-bands include an LLL (low-low-low) sub-band and an LLH (low-low-high) sub-band and the contexts are assigned as follows:

LLL and LLH Sub-bands				
<b>h</b>	<b>v</b>	<b>a</b>	<b>d</b>	<b>Context</b>
2	x	x	x	0
1	$\geq 1$	x	x	0
1	0	$\geq 1$	x	1
1	0	0	x	2
0	2	0	x	3
0	1	0	x	4
0	0	$\geq 1$	x	5
0	0	0	3	6
0	0	0	2	7
0	0	0	1	8
0	0	0	0	9

where "h" represents a number of immediate horizontal neighbors that are significant and  $0 < h < 2$ , "v" represents a number of immediate vertical neighbors that are significant and  $0 < v < 2$ , "a" represents a number of immediate temporal neighbors that are significant and  $0 < a < 2$ , and "d" represents a number of immediate diagonal neighbors that are significant and  $0 < d < 12$ .

12. A method as recited in claim 10, wherein the sub-bands include an LHH (low-high-high) sub-band and the contexts are assigned as follows:

LHH Sub-band			
<b>h</b>	<b>v+a</b>	<b>d</b>	<b>Context</b>
2	x	x	0
1	$\geq 3$	x	0
1	$\geq 1$	$\geq 4$	1
1	$\geq 1$	x	2
1	0	$\geq 4$	3
1	0	x	4
0	$\geq 3$	x	5
0	$\geq 1$	$\geq 4$	6
0	$\geq 1$	x	7
0	0	$\geq 4$	8
0	0	x	9

where "h" represents a number of immediate horizontal neighbors that are significant and  $0 < h < 2$ , "v" represents a number of immediate vertical neighbors that are significant and  $0 < v < 2$ , "a" represents a number of immediate temporal neighbors that are significant and  $0 < a < 2$ , and "d" represents a number of immediate diagonal neighbors that are significant and  $0 < d < 12$ .



1        16. A computer-readable medium comprising computer-executable  
2 instructions that, when executed by one or more processors, perform the method as  
3 recited in claim 1.

4  
5        17. A method comprising:  
6        estimating motion trajectories of pixels in a video object from frame to  
7 frame in a video sequence;  
8        performing a temporal wavelet transform on the corresponding pixels along  
9 the motion trajectories in a temporal direction to produce frames of temporal  
10 wavelet coefficients;  
11        performing a spatial wavelet transform on the frames of the temporal  
12 wavelet coefficients to produce multiple sub-bands of wavelet coefficients; and  
13        coding each sub-band of wavelet coefficients independently.

14  
15       18. A method as recited in claim 17, wherein the estimating comprises  
16 matching corresponding pixels in the video object from frame to frame in the  
17 video sequence.

18  
19       19. A method as recited in claim 17, wherein the temporal and spatial  
20 wavelet transforms comprise a shape-adaptive discrete wavelet transform.

21  
22       20. A method as recited in claim 17, wherein the performing a temporal  
23 wavelet transform comprises:

24       forming a pixel array containing pixels that continue from frame to frame in  
25 the video sequence;

09599160-062100

1 examining a pixel in a frame to determine whether the pixel is a terminating  
2 pixel that does not continue to a next frame;  
3 if the pixel is a terminating pixel, terminating the pixel array; and  
4 if the pixel is not a terminating pixel, adding the pixel to the pixel array.  
5

6 **21.** A method as recited in claim 20, further comprising transforming  
7 the pixels arrays to produce the frames of temporal wavelet coefficients.  
8

9 **22.** A method as recited in claim 17, wherein the coding comprises  
10 transposing selected sub-bands to reduce a number of sub-bands to be coded.  
11

12 **23.** A method as recited in claim 17, wherein the coding comprises:  
13 coding the wavelet coefficients in bit-planes; and  
14 allocating bits for the bit-planes according to a rate-distortion optimization.  
15

16 **24.** A method as recited in claim 17, further comprising truncating bits  
17 allocated to a bit-plane at a point on a rate-distortion curve that approximates a  
18 convex hull.  
19

20 **25.** A method as recited in claim 17, wherein the coding comprises  
21 coding the wavelet coefficients of each sub-band bit-plane by bit-plane using  
22 different coding primitives.  
23  
24  
25

1       **26.**   A method as recited in claim 25, wherein the coding primitives  
2 comprise:

3           zero coding to code new information about a wavelet coefficient that is not  
4 yet significant in a previous bit-plane; and

5           sign coding to code a sign of the wavelet coefficient once the wavelet  
6 coefficient is deemed significant.

7  
8       **27.**   A method as recited in claim 25, wherein the coding primitives  
9 comprise:

10          zero coding to code new information about a wavelet coefficient that is not  
11 yet significant in a previous bit-plane;

12          sign coding to code a sign of the wavelet coefficient once the wavelet  
13 coefficient is deemed significant; and

14          magnitude refinement to code new information of a wavelet coefficient that  
15 has already become significant in the previous bit-plane.

16  
17       **28.**   A method as recited in claim 17, wherein the coding produces  
18 multiple bitstreams for corresponding sub-bands of wavelet coefficients and  
19 further comprising constructing a multi-layer bitstream from the multiple  
20 bitstreams.

21  
22       **29.**   A method as recited in claim 17, wherein the coding comprises  
23 assigning contexts to the wavelet coefficients of each sub-band based on numbers  
24 of significant neighboring samples.



1       **30.**   A computer-readable medium comprising computer-executable  
2 instructions that, when executed by one or more processors, perform the method as  
3 recited in claim 17.

4  
5       **31.**   A method comprising:  
6       forming an array containing a current pixel in a current frame;  
7       determining whether a pixel corresponding to the current pixel resides in a  
8 next frame subsequent to the current frame;  
9       if the corresponding pixel exists, add the current pixel to the array; and  
10      if the corresponding pixel does not exist, terminate the pixel array with the  
11 current pixel.

12  
13      **32.**   A method as recited in claim 31, further comprising initially  
14 indicating all pixels as one state and changing the state of each pixel after said  
15 examining of the pixel.

16  
17      **33.**   A method as recited in claim 31, further comprising applying a  
18 wavelet transform to the pixel array to produce wavelet coefficients.

19  
20      **34.**   A method as recited in claim 31, further comprising:  
21      applying a wavelet transform to the pixel array to produce wavelet  
22 coefficients; and  
23      applying a spatial 2-D wavelet transform on the wavelet coefficients.  
24  
25

1        35. A computer-readable medium comprising computer-executable  
2 instructions that, when executed by one or more processors, perform the method as  
3 recited in claim 31.

4  
5        36. A method comprising:  
6 coding sub-bands of coefficients produced from transforming video frames  
7 in an independent manner such that one sub-band of coefficients is coded  
8 independently of another sub-band of coefficients; and  
9 constructing a bitstream from the independently coded sub-bands.

10  
11       37. A method as recited in claim 36, wherein the coding comprises  
12 transposing selected sub-bands prior to said coding.

13  
14       38. A method as recited in claim 36, wherein the coding comprises  
15 coding the coefficients of each sub-band bit-plane by bit-plane using different  
16 coding primitives.

17  
18       39. A method as recited in claim 38, wherein the coding primitives  
19 comprise:

20 zero coding to code new information about a coefficient that is not yet  
21 significant in a previous bit-plane; and

22 sign coding to code a sign of the coefficient once the coefficient is deemed  
23 significant.

1           **40.**   A method as recited in claim 38, wherein the coding primitives  
2 comprise:

3           zero coding to code new information about a coefficient that is not yet  
4 significant in a previous bit-plane;

5           sign coding to code a sign of the coefficient once the coefficient is deemed  
6 significant; and

7           magnitude refinement to code new information of a coefficient that has  
8 already become significant in the previous bit-plane.

9  
10          **41.**   A method as recited in claim 36, wherein the coding comprises  
11 assigning contexts to the coefficients of each sub-band based on numbers of  
12 significant neighboring samples.

13  
14          **42.**   A method as recited in claim 41, wherein the sub-bands include an  
15 LLL (low-low-low) sub-band and an LLH (low-low-high) sub-band and the  
16 contexts are assigned as follows:

LLL and LLH Sub-bands				
<b>h</b>	<b>v</b>	<b>a</b>	<b>d</b>	<b>Context</b>
2	x	x	x	0
1	$\geq 1$	x	x	0
1	0	$\geq 1$	x	1
1	0	0	x	2
0	2	0	x	3
0	1	0	x	4
0	0	$\geq 1$	x	5
0	0	0	3	6
0	0	0	2	7
0	0	0	1	8
0	0	0	0	9

where “h” represents a number of immediate horizontal neighbors that are significant and  $0 < h < 2$ , “v” represents a number of immediate vertical neighbors that are significant and  $0 < v < 2$ , “a” represents a number of immediate temporal neighbors that are significant and  $0 < a < 2$ , and “d” represents a number of immediate diagonal neighbors that are significant and  $0 < d < 12$ .

43. A method as recited in claim 41, wherein the sub-bands include an LHH (low-high-high) sub-band and the contexts are assigned as follows:

LHH Sub-band			
<b>h</b>	<b>v+a</b>	<b>d</b>	<b>Context</b>
2	x	x	0
1	$\geq 3$	x	0
1	$\geq 1$	$\geq 4$	1
1	$\geq 1$	x	2
1	0	$\geq 4$	3
1	0	x	4
0	$\geq 3$	x	5
0	$\geq 1$	$\geq 4$	6
0	$\geq 1$	x	7
0	0	$\geq 4$	8
0	0	x	9

where “h” represents a number of immediate horizontal neighbors that are significant and  $0 < h < 2$ , “v” represents a number of immediate vertical neighbors that are significant and  $0 < v < 2$ , “a” represents a number of immediate temporal neighbors that are significant and  $0 < a < 2$ , and “d” represents a number of immediate diagonal neighbors that are significant and  $0 < d < 12$ .

1           44.     A method as recited in claim 41, wherein the sub-bands include an  
2 HHH (high-high-high) sub-band and the contexts are assigned as follows:

3

d	h+v+a	Context
$\geq 6$	x	0
$\geq 4$	$\geq 3$	1
$\geq 4$	x	2
$\geq 2$	$\geq 4$	3
$\geq 2$	$\geq 2$	4
$\geq 2$	x	5
$\geq 0$	$\geq 4$	6
$\geq 0$	$\geq 2$	7
$\geq 0$	1	8
$\geq 0$	0	9

10

11

12 where “h” represents a number of immediate horizontal neighbors that are  
13 significant and  $0 < h < 2$ , “v” represents a number of immediate vertical neighbors  
14 that are significant and  $0 < v < 2$ , “a” represents a number of immediate temporal  
15 neighbors that are significant and  $0 < a < 2$ , and “d” represents a number of  
16 immediate diagonal neighbors that are significant and  $0 < d < 12$ .

17

18           45.     A method as recited in claim 36, wherein the constructing comprises  
19 forming multiple bit-planes and truncating a number of bits in each bit-plane  
20 according to a rate-distortion curve.

21

22           46.     A computer-readable medium comprising computer-executable  
23 instructions that, when executed by one or more processors, perform the method as  
24 recited in claim 36.

25

1       **47.**    A method comprising:  
2       transforming a video sequence to produce a set of sub-bands;  
3       transposing selected sub-bands to produce a reduced set of sub-bands that  
4       are fewer than the set of sub-bands; and  
5       coding the reduced set of sub-bands.

6  
7       **48.**    A method as recited in claim 47, wherein the coding comprises  
8       coding each sub-band independently.

9  
10       **49.**   A method as recited in claim 47, wherein the coding comprises  
11       assigning contexts to the coefficients of each sub-band based on numbers of  
12       significant neighboring samples.

13  
14       **50.**   A computer-readable medium comprising computer-executable  
15       instructions that, when executed by one or more processors, perform the method as  
16       recited in claim 47.

17  
18       **51.**   A method for coding coefficients indicative of transformed video in  
19       multiple bit-planes, comprising:

20       conducting coding passes through a bit-plane to code significant samples  
21       separately from insignificant samples; and  
22       repeating said conducting for each bit-plane.

a significant propagation pass to code samples that are not yet significant;

a magnitude refinement pass to code samples that are already deemed significant; and

a normalization pass to code any samples not coded in the significant propagation pass and the magnitude refinement passes.

**53.** A computer-readable medium comprising computer-executable instructions that, when executed by one or more processors, perform the method as recited in claim 51.

**54.** A video encoder comprising:

- a wavelet transformer to transform frames in a video sequence into multiple sub-bands of coefficients, the wavelet transform using motion information of video objects in the frames; and
- a coder to code the coefficients of each sub-band independently.

**55.** A video encoder as recited in claim 54, wherein the wavelet transformer applies a spatial-adaptive discrete wavelet transform to the frames.

56. A video encoder as recited in claim 54, wherein the wavelet transformer applies a temporal wavelet transform on pixels in the frames taken along motion trajectories in a temporal direction from frame to frame.





62. A video encoder as recited in claim 61, wherein the sub-bands include an LLL (low-low-low) sub-band and an LLH (low-low-high) sub-band and the coder employs a zero coding primitive to code new information about a coefficient that is not yet significant in a previous bit-plane by assigning the contexts as follows:

LLL and LLH Sub-bands				
<b>h</b>	<b>v</b>	<b>a</b>	<b>d</b>	<b>Context</b>
2	x	x	x	0
1	$\geq 1$	x	x	0
1	0	$\geq 1$	x	1
1	0	0	x	2
0	2	0	x	3
0	1	0	x	4
0	0	$\geq 1$	x	5
0	0	0	3	6
0	0	0	2	7
0	0	0	1	8
0	0	0	0	9

where "h" represents a number of immediate horizontal neighbors that are significant and  $0 < h < 2$ , "v" represents a number of immediate vertical neighbors that are significant and  $0 < v < 2$ , "a" represents a number of immediate temporal neighbors that are significant and  $0 < a < 2$ , and "d" represents a number of immediate diagonal neighbors that are significant and  $0 < d < 12$ .

63. A video encoder as recited in claim 61, wherein the sub-bands include an LHH (low-high-high) sub-band and the coder employs a zero coding primitive to code new information about a coefficient that is not yet significant in a previous bit-plane by assigning the contexts as follows:

LHH Sub-band			
h	v+a	d	Context
2	x	x	0
1	$\geq 3$	x	0
1	$\geq 1$	$\geq 4$	1
1	$\geq 1$	x	2
1	0	$\geq 4$	3
1	0	x	4
0	$\geq 3$	x	5
0	$\geq 1$	$\geq 4$	6
0	$\geq 1$	x	7
0	0	$\geq 4$	8
0	0	x	9

where "h" represents a number of immediate horizontal neighbors that are significant and  $0 < h < 2$ , "v" represents a number of immediate vertical neighbors that are significant and  $0 < v < 2$ , "a" represents a number of immediate temporal neighbors that are significant and  $0 < a < 2$ , and "d" represents a number of immediate diagonal neighbors that are significant and  $0 < d < 12$ .

64. A video encoder as recited in claim 61, wherein the sub-bands include an HHH (high-high-high) sub-band and the coder employs a zero coding primitive to code new information about a coefficient that is not yet significant in a previous bit-plane by assigning the contexts as follows:

d	h+v+a	Context
$\geq 6$	x	0
$\geq 4$	$\geq 3$	1
$\geq 4$	x	2
$\geq 2$	$\geq 4$	3
$\geq 2$	$\geq 2$	4
$\geq 2$	x	5
$\geq 0$	$\geq 4$	6
$\geq 0$	$\geq 2$	7
$\geq 0$	1	8
$\geq 0$	0	9

where "h" represents a number of immediate horizontal neighbors that are significant and  $0 < h < 2$ , "v" represents a number of immediate vertical neighbors that are significant and  $0 < v < 2$ , "a" represents a number of immediate temporal neighbors that are significant and  $0 < a < 2$ , and "d" represents a number of immediate diagonal neighbors that are significant and  $0 < d < 12$ .

65. A video encoder as recited in claim 61, wherein the coder employs a sign coding primitive to code a sign of the coefficient once the coefficient is deemed significant by assigning the contexts as follows:

h=-1			
v	a	$\hat{x}$	Context
-1	-1	0	0
-1	0	0	1
-1	1	0	2
0	-1	0	3
0	0	0	4
0	1	0	5
1	-1	0	6
1	0	0	7
1	1	0	8

H=0			
v	a	$\hat{x}$	Context
-1	-1	0	9
-1	0	0	10
-1	1	0	11
0	-1	0	12
0	0	0	13
0	1	1	12
1	-1	1	11
1	0	1	10
1	1	1	9

h=1			
v	a	$\hat{x}$	Context
-1	-1	1	8
-1	0	1	7
-1	1	1	6
0	-1	1	5
0	0	1	4
0	1	1	3
1	-1	1	2
1	0	1	1
1	1	1	0

where “h” represents a number of immediate horizontal neighbors that are significant and  $0 < h < 2$ , “v” represents a number of immediate vertical neighbors that are significant and  $0 < v < 2$ , “a” represents a number of immediate temporal neighbors that are significant and  $0 < a < 2$ , and  $\hat{x}$  is a sign symbol prediction in a given context.

007290" 09T66560

1        66.    A video encoder as recited in claim 54, wherein the coder truncates  
2 a number of coding bits according to rate-distortion curves.

3  
4        67.    An operating system embodied on a computer-readable medium  
5 comprising a video encoder as recited in claim 54.

6  
7        68.    A video encoder comprising:  
8        means for estimating motion trajectories of pixels in a video object from  
9 frame to frame in a video sequence;  
10       means for performing a temporal wavelet transform on the corresponding  
11 pixels along the motion trajectories in a temporal direction to produce frames of  
12 temporal wavelet coefficients;  
13       means for performing a spatial wavelet transform on the frames of the  
14 temporal wavelet coefficients to produce multiple sub-bands of wavelet  
15 coefficients; and  
16       means for coding each sub-band of wavelet coefficients independently.

17  
18       69.    A video encoder as recited in claim 68, wherein the estimating  
19 means comprises means for matching corresponding pixels in the video object  
20 from frame to frame in the video sequence.

21  
22       70.    A video encoder as recited in claim 68, wherein the temporal and  
23 spatial wavelet transforms comprise a shape-adaptive discrete wavelet transform.  
24  
25



007290" 09T66560

1       75.    A video encoder as recited in claim 68, wherein the coding means  
2 comprises means for coding the wavelet coefficients of each sub-band bit-plane by  
3 bit-plane using different coding primitives.

4  
5       76.    A video encoder as recited in claim 68, wherein the coding means  
6 produces multiple bitstreams for corresponding sub-bands of wavelet coefficients  
7 and further comprising means for constructing a multi-layer bitstream from the  
8 multiple bitstreams.

9  
10       77.   A video encoder as recited in claim 68, wherein the coding means  
11 comprises means for assigning contexts to the wavelet coefficients of each sub-  
12 band based on numbers of significant neighboring samples.

13  
14       78.   A computer-readable medium comprising computer-executable  
15 instructions that, when executed on a processor, direct a device to:

16       code sub-bands of coefficients produced from transforming video frames in  
17 an independent manner such that one sub-band of coefficients is coded  
18 independently of another sub-band of coefficients; and

19       construct a bitstream from the independently coded sub-bands.

20  
21       79.   A computer-readable medium as recited in claim 78, further  
22 comprising computer-executable instructions that, when executed on a processor,  
23 direct a device to transpose selected sub-bands prior to said coding.



001290" 09T6560

1       **80.**   A computer-readable medium as recited in claim 78, further  
2 comprising computer-executable instructions that, when executed on a processor,  
3 direct a device to code the coefficients of each sub-band bit-plane by bit-plane  
4 using different coding primitives.

5  
6       **81.**   A computer-readable medium as recited in claim 78, further  
7 comprising computer-executable instructions that, when executed on a processor,  
8 direct a device to assign contexts to the coefficients of each sub-band based on  
9 numbers of significant neighboring samples.

10  
11       **82.**   A computer-readable medium embodying an encoded video signal  
12 constructed as a result of a process comprising:

13       transforming frames in a video sequence using a wavelet transform to  
14 produce multiple sub-bands of coefficients;

15       coding the coefficients of each sub-band independently to produce multiple  
16 bitstreams; and

17       forming a bitstream from the multiple bitstreams.  
18  
19  
20  
21  
22  
23  
24  
25